

KOREAN PATENT ABSTRACT (KR)

PUBLICATION

(51) IPC Code: H03M 1/12

(11) Publication No.: P1998-079475

(43) Publication Date: 25 November 1998

(21) Application No.: 10-1997-061298

(22) Application Date: 19 November 1997

(71) Applicant:

Samsung Electronics Co., Ltd. Jong-yong Yun
416 Maetan-dong, Paldal-gu, Suwon-City, 442-742 Korea

(72) Inventor:

Seong-hui Park
Hanil Apt. ga-506, 1641-1 Seocho 1-dong, Seocho-gu, Seoul, 137-071 Korea

(54) Title of the Invention:

Method and apparatus for encoding/decoding bit rate scalable audio signal

Abstract:

Provided are a method and an apparatus for encoding/decoding bit rate scalable audio signal. To encode an audio signal into a hierarchical data stream configured of a base layer and a predetermined number of enhancement layers, an input audio signal is processed to be quantized by predetermined encoding bands. Quantized data corresponding to the basic layer is encoded within a predetermined layer size. Quantized data belonging to the enhancement layer following the encoded layer and quantized data belonging to the encoded layer that is not encoded are encoded within a predetermined layer size. The layer encoding is performed for all layers. The encoding process includes the first step of representing quantized data corresponding to layers to be encoded as the same number of digits and the second step of encoding the quantized data represented as the same number of digits in sequence from more significant digit sequence to less significant digit sequence. Therefore, the complexity of an encoder/decoder decreases. Also, the bit rate and complexity of an encoder/decoder can be adjusted according to the performance of a decoder, the bandwidth/complexity of a transmission line, or user's request.

(19) 대한민국특허청(KR)

(12) 공개특허공보(A)

(51) Int. Cl.
H03M 1/12

(11) 공개번호
(43) 공개일자

특1998-079475
1998년11월25일

(21) 출원번호	특1997-061298
(22) 출원일자	1997년11월19일
(30) 우선권주장	101997012232 1997년04월02일 대한민국(KR)
(71) 출원인	삼성전자 주식회사, 윤종용 대한민국 442-742 경기도 수원시 팔달구 매탄동 416번지
(72) 발명자	박성희 대한민국 137-071 서울특별시 서초구 서초1동 1641-1 한일아파트 가동 506호
(74) 대리인	이영필 권석홍 이상용
(77) 심사청구	있음
(54) 출원명	비트율 조절이 가능한 오디오 부호화/복호화 방법 및 장치

요약

본 발명은 비트율 조절이 가능한 오디오 부호화기/복호화 방법 및 장치에 관한 것으로서, 오디오 부호화 방법은 오디오 신호를 base layer과 소정 개수의 enhancement layer으로 이루어지는 계층적인 데이터스트림으로 부호화하는 방법에 있어서 입력 오디오 신호를 신호처리하여 소정의 부호화 대역별로 양자화하는 단계, 기본계층에 상응하는 양자화된 데이터들을 소정의 layer 크기 내에서 부호화하는 기본계층 부호화단계, 부호화된 계층의 다음 enhancement 계층에 속하는 양자화된 데이터 및 부호화된 계층에 속하는 부호화되지 않고 남아있는 양자화된 데이터들을 소정의 계층 크기 내에서 부호화하는 부호화단계 및 그 부호화단계를 모든 계층에 대해 수행하는 단계를 포함함을 특징으로 하고, 그 부호화는 부호화하고자 하는 해당 계층에 상응하는 양자화된 데이터를 소정의 동일 개수의 디지털로 표현하는 제1단계 및 동일 개수의 디지털로 표현된 양자화된 데이터들의 중요도가 가장 높은 디지털들로 이루어지는 최상위 디지털 시퀀스부터 중요도가 낮은 디지털 시퀀스 순으로 소정의 부호화 방법으로 부호화하는 제2단계로 이루어진다.

본 발명에 의하면, 전송선로의 상태나 사용자의 다양한 요구도 수용할 수 있으며, 부호화기 및 복호화기의 복잡도가 낮다. 또한 복호화기의 성능, 전송선로의 대역폭/혼잡도 또는 사용자의 요구에 따라서 비트율이나 복잡성 조절이 가능하다.

대표도

도2

명세서

도면의 간단한 설명

도 1은 단순한 개념의 스케일 조절가능한 Codec의 구성을 블록도로 도시한 것이다.

도 2는 본 발명에 의한 부호화장치의 구조를 블록도로 도시한 것이다.

도 3은 본 발명에 의한 비트스트림의 구조를 도시한 것이다.

도 4는 본 발명에 의한 복호화장치의 구조를 블록도로 도시한 것이다.

도 5는 긴 블록의 경우 양자화 대역과 부호화 대역을 도시한 것이다.

도 6은 짧은 블록의 경우 양자화 대역과 부호화 대역을 도시한 것이다.

도 7은 비트율에 따른 1프레임의 비트 수를 도시한 것이다.

도 8은 비트율에 따른 대역제한을 도시한 것이다.

도 9는 8개의 짧은 블록을 나눈 3개의 영역의 형태를 도시한 것이다.

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 오디오 부호화/복호화에 관한 것으로서, 특히 하나의 비트스트림내에 1개의 비트율에 대한 것만으로 구성하는 것이 아니라 기본 계층(Base Layer)를 기반으로 여러 상위 계층(Enhancement Layer)의 데이터를 함께 표현하는 계층구조의 비트스트림을 부호화하는, 비트율 조절이 가능한 오디오 부호화/복호화 방법 및 장치에 관한 것이다.

본 발명은 기존의 부호화 방식들과 같이 정해진 비트율에서 최상의 성능을 나타내는 부호화 효율 측면을 강조하는 방식이 아닌 멀티미디어 시대에 적합하게 부호화방식으로서, 스케일 조절이 가능한(scalable) 부호화/복호화 기법에 관한 것이다. 그 부호화장치는 기본계층의 비트율 이상의 비트율 중에서 사용자가 지정한 비트율로 비트스트림을 제작하고, 그 복호화장치는 상기 부호화된 비트율 이내의 비트율 중의 하나로 복원할 수 있다.

정보를 포함하고 있는 파형은 원래 진폭에 있어서 연속적이고 시간상으로도 연속적인 아날로그(Analog) 신호이다. 따라서, 파형을 이산 신호로 표현하기 위해서 A/D(Analog-to-Digital) 변환이 필요하다. A/D 변환을 하기 위해서 두 가지의 과정을 필요로 한다. 하나는 시간상의 연속 신호를 이산 신호로 바꾸어 주는 표본화(Sampling)과정이고, 다른 하나는 가능한 진폭의 수를 유한한 값으로 제한하기 위한 진폭 양자화 과정이다. 즉, 진폭의 양자화는 시간 n 에서 입력 진폭 $x(n)$ 을 가능한 진폭의 유한한 집합 중의 한 원소인 $y(n)$ 으로 변환해 주는 과정이다.

오디오 신호의 저장/복원 방식도 최근 디지털 신호처리 기술의 발달에 의해 기존의 아날로그 신호를 표본화와 양자화 과정을 거쳐 디지털 신호인 PCM(Pulse Code Modulation) 데이터로 변환하여 CD(Compact Disc)와 DAT(Digital Audio Tape)와 같은 기록/저장 매체에 신호를 저장해 둔 뒤 사용자가 필요시에 저장된 신호를 다시 재생해서 들을 수 있는 기술이 개발되어 일반인들에게 보편화되어 사용되고 있다. 이런 디지털 방식에 의한 저장/복원 방식은 LP(Long-Play Record)와 Tape와 같은 아날로그 방식에 비해 음질의 향상과 저장 기간에 따른 열화를 극복하였으나 디지털 데이터의 크기가 많아 저장 및 전송에 문제를 보였다.

이러한 문제를 해결하기 위해 디지털 음성 신호를 압축하기 위해 개발된 DPCM(Differential Pulse Code Modulation)이나 ADPCM(Adaptive Differential Pulse Code Modulation)등의 방법을 사용하여 데이터의 양을 줄이려는 노력이 있었으나 신호의 종류에 따라 효율성이 크게 차이가 나는 단점을 보였다. 최근 ISO (International Standard Organization)에 의해 표준화 작업이 이루어진 MPEG/audio(Moving Pictures Expert Group) 기법이나 Dolby에 의해 개발된 AC-2/AC-3 기법에서는 인간의 심리음향 모델(Psychoacoustic Model)을 이용하여 데이터의 양을 줄이는 방법을 사용했다. 이러한 방법은 신호의 특성에 관계없이 효율적으로 데이터의 양을 줄이는데 크게 기여하였다.

MPEG-1/audio, MPEG-2/audio 나 AC-2/AC-3등과 같은 기존의 오디오 신호 압축 기법에서는 시간영역의 신호를 일정한 크기의 블럭으로 묶어서 주파수 영역의 신호로 변환한다. 그리고 이 변환된 신호를 인간의 심리음향 모델(Psychoacoustic Model)을 이용하여 스칼라 양자화한다. 이런 양자화 기법은 단순하지만 입력 샘플이 통계적으로 독립적이라고 할지라도 최적은 아니다. 물론 입력 샘플이 통계적으로 종속적이라면 더욱 불충분하다. 이런 문제점 때문에 엔트로피(Entropy) 부호화와 같은 무손실 부호화나 어떤 종류의 적응 양자화를 포함하여 부호화를 수행한다. 따라서, 단순한 PCM 데이터만을 저장하던 방식보다는 상당히 복잡한 과정을 거치고 비트스트림은 양자화된 PCM 데이터뿐만 아니라 신호를 압축하기 위한 추가적인 정보들로 구성되어 있다.

MPEG/audio 표준이나 AC-2/AC-3 방식은 기존의 디지털 부호화에 비해 1/6내지 1/8로 줄어든 64Kbps - 384Kbps 의 비트율로 콤팩트디스크(Compact Disc) 음질과 거의 같은 정도의 음질을 제공한다. 이런 이유때문에, MPEG/audio 표준은 DAB(Digital Audio Broadcasting), 인터넷 폰(internet phone), AOD(Audio on Demand) 와 멀티미디어 시스템과 같은 오디오 신호의 저장과 전송에 중요한 역할을 할 것이다.

이런 기존의 기법들은 부호화기에서 고정된 비트율이 주어지고 주어진 비트율에 최적의 상태를 찾아 양자화와 부호화 과정을 거치기 때문에 고정된 비트율을 사용할 경우에는 상당히 좋은 방안을 제시한다. 하지만 멀티미디어 시대가 도래함과 더불어 기존의 낮은 비트율 부호화뿐 아니라 여러가지 가능성을 가지고 있는 부호화기/복호화기에 대한 요구가 많아지고 있다. 그 요구 중의 하나가 비트율의 크기 조절이 가능한 오디오 부호화/복호화기이다. 비트율 조절이 가능한 오디오 부호화기는 높은 비트율로 부호화된 비트스트림을 낮은 비트율의 비트스트림으로 만들 수도 있고, 그 중 일부의 비트스트림만을 가지고 복원을 할 수 있게 한다. 이렇게 함으로써 네트워크에서 과부하가 걸리거나 복호화기에서의 성능이 좋지 않을 경우 또는 사용자의 요구에 의해 비트율이 낮아질 경우에는 비트스트림의 일부만을 가지고 비트율이 낮아진만큼 성능의 열화를 보이지만 어느 정도의 성능으로 사용자에게 신호를 복원해 줄 수 있다.

기존의 오디오 부호화 기법들은 부호화기에 고정된 비트율이 주어지고 상기 주어진 비트율에 대해 최적의 상태를 찾아 양자화와 부호화 과정을 거쳐 주어진 비트율에 맞춰 비트스트림을 만들어 낸다. 따라서, 1개의 비트스트림 내에 1개의 비트율에 대한 정보만을 담고 있다. 즉, 비트스트림 헤더(header)에 비트율에 대한 정보를 담고 고정적인 비트율을 사용한다. 이런 방법의 장점은 고정된 특정 비트율만을 사용할 경우에 상기 특정 비트율에서 나타낼 수 있는 가장 좋은 방안을 사용할 수 있다는 것이다. 예를 들면, 한 비트스트림이 96Kbps의 비트율로 부호화기에서 만들어졌다면 복호화기에서는 96Kbps에 상응하는 복호화기로 복원한다면 96Kbps의 비트율에서는 최상의 품질로 복원된 소리를 들을 수 있다.

이러한 방식의 비트스트림 구성은 다른 비트율을 고려하지 않은 채 비트스트림이 구성되어 있기 때문에 비트스트림의 순서에 대한 고려보다는 주어진 비트율에 알맞은 크기로 비트스트림이 구성되어 있다. 실제로 이렇게 구성된 비트스트림이 네트워크를 통해서 전송될 경우 이 비트스트림은 몇 개의 슬롯(slot)으로 쪼개어 보내어 지게 된다. 전송선로에 과부하가 걸린다든지 전송선로의 대역폭이 좁아서 수신단에 송신단에서 보낸 슬롯 전체가 도착하지 못하고 일부만이 도착하게 될 경우 올바른 데이터를 재생할 수 없다. 또한 비트스트림의 순서가 중요도에 따라서 구성되어 있지 않기 때문에 전체 비트스트림이 아닌 일부 비트스트림만을 가지고 복원할 경우 상당히 귀에 거슬리는 소리를 재생하게 된다.

하나의 방송국에서 비트스트림을 제작하여 여러 사용자에게 전송하는 경우를 고려해 보기로 한다. 각 비트스트림을 수신하는 사용자마다 서로 다른 비트율을 요구하거나 서로 다른 성능의 복호화기를 가지고 있을 경우, 고정비트율만을 지원하는 비트스트림만을 방송국에서 송신한다면 이러한 요구사항을 만족시키기 위해 각 사용자마다 각각의 비트스트림을 전송해야 할 것이다. 이것은 전송하는데 상당한 비용이 들고 또한 방송국에서는 각 사용자의 요구에 맞는 비트율의 비트스트림을 제작하거나 저장하고 있어야 하기 때문에 여기에 들어가야 할 비용도 상당히 많다.

이러한 경우 오디오의 비트스트림이 여러 가지 계층의 비트율로 구성이 되어있다면 주어진 환경이나 사용자의 요구사항에 대해 적절하게 대처할 수 있을 것이다. 이런 구조를 가지기 위한 가장 단순한 방법은 도 1과 같이 낮은 계층에 대한 부호화를 행한후 다시 복호화하여 얻은 신호와 원래신호와 차이를 다시 다음 계층의 부호화기 입력으로 사용하여 처리하는 방식이다. 즉 먼저 기본 계층의 비트율에 대해 부호화를 해서 비트스트림을 생성하고 난 후 원래 신호와 부호화된 신호의 차이 신호를 가지고 다시 부호화기를 해서 다음 계층의 비트스트림을 생성하는 과정을 반복하게 된다. 따라서 이러한 방법은 부호화기의 복잡도가 높아지고 또한 복호화기에서도 같은 과정을 역으로 반복해야 복원된 신호를 얻을 수 있기 때문에 복호화기의 복잡성도 증가하게 된다. 따라서 계층의 수가 늘어나면 늘어날 수록 부호화기와 복호화기의 복잡성이 비례해서 늘어나난다.

발명이 이루고자 하는 기술적 과제

본 발명이 이루고자 하는 기술적 과제는, 여러 계층의 비트율에 대한 데이터를 하나의 비트스트림내에 표현함으로써 전송선로의 상태, 복호화기의 성능 또는 사용자의 요구사항에 따라 비트스트림의 크기가 유동적으로 바뀔 수도 있고 복호화기의 복잡성도 바뀔 수도 있는, 비트율 조절이 가능한 오디오 부호화기/복호화기 및 그 방법을 제공함에 있다.

발명의 구성 및 작용

상기의 기술적 과제를 만족하기 위한, 본 발명에 의한 비트율 조절이 가능한 오디오 부호화장치는, 오디오 신호를 기본계층(base layer)과 소정 개수의 향상계층(enhancement layer)으로 이루어지는 계층적인 데이터스트림으로 부호화하는 방법에 있어서, 입력 오디오 신호를 신호처리하여 소정의 부호화 대역별로 양자화하는 양자화처리단계; 상기 기본계층에 상응하는 양자화된 데이터들을 소정의 layer 크기 내에서 부호화하는 기본계층 부호화단계; 상기 부호화된 계층의 다음 enhancement 계층에 속하는 양자화된 데이터 및 상기 부호화된 계층에 속하는 부호화되지 않고 남아있는 양자화된 데이터들을 소정의 계층 크기 내에서 부호화하는 계층부호화 단계; 및 상기 계층부호화단계를 모든 계층에 대해 수행하는 순차부호화단계를 포함함을 특징으로 한다. 상기 기본계층부호화단계, 계층부호화단계 및 순차부호화단계의 부호화는 부호화하고자 하는 해당 계층에 상응하는 양자화된 데이터를 소정의 동일 개수의 디지털로 표현하는 제1단계; 및 상기 동일 개수의 디지털로 표현된 양자화된 데이터들의 중요도가 가장 높은 디지털들로 이루어지는 최상위 디지털 시퀀스부터 중요도가 낮은 디지털 시퀀스 순으로 소정의 부호화 방법으로 부호화하는 제2단계로 이루어진다.

상기 제1단계 및 제2단계의 수행은, 낮은 주파수성분부터 높은 주파수성분 순서대로 수행함이 바람직하다.

상기 기본계층부호화단계, 계층부호화단계 및 순차부호화단계의 부호화는 적어도 스텝크기(step size) 및 각 대역에 할당된 양자화비트수를 포함하는 부가정보를 소정의 부호화 방법을 사용하여 부호화함이 바람직하다.

상기 제1단계 및 제2단계의 디지털트는 비트이고, 상기 제2단계의 부호화는 상기 비트 시퀀스를 구성하고 있는 비트들을 소정 개수의 비트 단위로 묶어서 부호화함이 바람직하다.

상기 소정의 부호화방법은 무손실 부호화이며, 허프만 부호화(Huffman coding) 또는 산술 부호화(Arithmetic coding) 임이 바람직하다.

한편, 상기 양자화된 데이터가 부호(sign)데이터와 크기(magnitude)데이터로 이루어질 때, 상기 제2단계는 상기 동일 개수의 디지털로 표현된 양자화된 데이터들의 크기 데이터들의 중요도가 가장 높은 최상위 디지털(Most Significant Digit)들로 이루어지는 최상위 디지털 시퀀스(sequence)를 소정의 부호화방법에 의해 부호화하는 단계; 상기 부호화된 최상위 디지털 시퀀스 중 영이 아닌 데이터에 해당하는 부호 데이터들을 부호화하는 단계; 상기 디지털 데이터의 부호화되지 않은 크기데이터 중 중요도가 가장 높은 디지털 시퀀스를 소정의 부호화 방법에 의해 부호화하는 크기부호화단계; 상기 제3단계에서 부호화된 디지털 시퀀스 중 영이 아닌 크기 데이터에 해당하는 부호데이터들 중 부호화되지 않은 부호(sign)데이터를 부호화하는 부호(sign)부호화단계; 및 상기 크기부호화단계 및 부호부호화단계를 상기 디지털 데이터의 각 디지털트에 대해 수행하는 단계를 포함함이 바람직하다.

상기 제1단계는 상기 디지털 데이터들을 동일 개수의 비트를 가진 이진 데이터로 표현하는 단계이고, 상기 디지털트는 비트임이 바람직하다.

상기 부호화는 상기 크기 데이터 및 부호(sign) 데이터에 대한 각 비트 시퀀스를 구성하고 있는 비트들을 소정 개수의 비트 단위로 묶어서 부호화한다.

상기 양자화처리단계는, 시간 영역의 입력 오디오 신호를 주파수 영역의 신호로 변환하는 단계; 상기 시간/주파수 맵핑에 의해 주파수 성분으로 변환된 입력 오디오 신호들을 소정의 대역(subband) 신호들로 묶고, 상기 각 대역(subband)에서의 마스킹 문턱치(masking threshold)를 계산하는 단계; 및 각 대역의 양자화 잡음이 마스킹 문턱치보다 작아지도록 소정의 부호화 대역별로 양자화하는 단계로 이루어짐이 바람직하다.

상기 기술적 과제를 달성하기 위한 오디오 신호를 소정 개수의 다른 비트율로 이루어지는 계층적인 비트율을 갖도록 부호화하는 장치는, 입력 오디오 신호를 신호처리하여 소정의 부호화 대역별로 양자화하는 양자화처리부; 및 기본계층에 상응하는 부가정보(side information) 및 양자화된 데이터들을 부호화하고, 상기 기본계층에 대한 부호화가 끝나면 그 다음 계층에 대한 부가정보 및 양자화된 데이터들을 부호하여 이를 모든 계층에 대해 수행하여 비트스트림을 형성하는 비트패킹부를 포함함을 특징으로 하고, 상기 비트패킹부의 각 계층에서의 양자화된 데이터의 부호화는 상기 양자화된 데이터 각각을 소정의 동일 개수의 비트로 이루어지는 이진 데이터로 표현하여 비트단위로 나누고, 상기 분할된 비트들로부터 중요도가 가장 높은 최상위 비트로 이루어지는 최상위 비트시퀀스부터 최하위 비트시퀀스 순서대로 소정의 부호화 방법에 의해 부호화한다.

그리고 상기 디지털 데이터가 부호(sign) 데이터와 크기(magnitude) 데이터로 이루어질 때 상기 비트패킹부는, 상기 비트분할된 데이터들 중 중요도가 같은 동일 순위의 비트들에 대한 크기데이터를 모아서 부호화하고, 상기 부호화된 크기데이터 중 영이 아닌 크기데이터에 해당하는 부호(sign)데이터 중 부호화되지 않은 부호데이터를 부호화하고, 상기 크기데이터와 부호데이터의 부호화를 최상위 비트부터 하위비트로 내려가며 순차적으로 수행함이 바람직하다.

상기 비트패킹부는, 중요도에 따라 비트들을 모아서 부호화할 때, 소정 개수의 비트단위로 묶어서 부호화하고, 낮은 주파수 성분부터 높은 주파수 성분 순서로 수행함이 바람직하다.

상기 기술적 과제를 달성하기 위한 본 발명에 의한, 계층적인 비트율을 갖도록 부호화된 오디오 데이터를 복호화하는 방법은, 계층적 구조를 가진 데이터스트림에서 상기 계층이 생성된 순서에 따라 적어도 양자화 스텝 크기 및 각 대역에 할당된 양자화비트수를 포함하는 부가정보 및 양자화된 데이터를 복호화하고, 상기 데이터스트림을 구성하고 있는 비트들의 중요도를 분석하여 중요도가 높은 디지털부터 중요도가 낮은 디지털로 내려가면서 복호화하는 복호화단계; 상기 복호화된 양자화 스텝 크기와 양자화된 데이터들을 원래 크기의 신호로 복원하는 단계; 및 상기 역양자화된 주파수 영역의 오디오 신호를 시간 영역의 신호로 변환하는 단계를 포함함이 바람직하다.

상기 복호화단계의 디지털트는 비트이고, 상기 복호화단계의 데이터 스트림은 비트스트림임이 바람직하다.

상기 복호화단계의 중요도에 따른 복호화는 소정의 개수의 비트로 이루어진 벡터단위로 복호화함이 바람직하다.

상기 양자화된 데이터가 부호(sign)데이터 및 크기(magnitude)데이터로 이루어질 때, 상기 복호화단계의 복호화는 계층적 구조를 가진 데이터 스트림에서 상기 계층이 생성된 순서에 따라 적어도 양자화 스텝의 크기 및 각 대역에 할당된 양자화비트수를 포함하는 부가정보 및 양자화된 크기 데이터를 복호화하고, 상기 데이터스트림을 구성하고 있는 비트들의 중요도를 분석하여 중요도가 높은 디지털부터 중요도가 낮은 디지털로 내려가면서 복호화하는 단계; 및 상기 양자화된 데이터의 부호(sign)데이터를 복호화하여 이를 상기 복호화된 크기 데이터와 결합하는 단계로 이루어짐을 특징으로 한다.

상기 복호화단계에서 복호화는 산술복호화 방법 또는 허프만 복호화방법에 의해 복호화함을 특징으로 한다.

상기 다른 기술적 과제를 해결하기 위한 계층적인 비트율을 갖도록 부호화된 오디오 데이터를 복호화하는 장치는, 계층적 구조를 가진 비트스트림에서 상기 계층이 생성된 순서에 따라 적어도 양자화 스텝의 크기 및 각 대역에 할당된 양자화비트수를 포함하는 부가정보 및 양자화된 데이터를 복호화하고, 상기 비트스트림을 구성하고 있는 비트들의 중요도를 분석하여 중요도가 높은 비트부터 중요도가 낮은 비트로 내려가면서 상기 계층별로 복호화하는 비트스트림분석부; 상기 비트스트림분석부에서 복호화된 양자화 스텝 크기와 양자화된 데이터들을 원래 크기의 신호로 복원하는 역양자화부; 및 상기 역양자화부에서 역양자화된 주파수 영역의 오디오 신호를 시간 영역의 신호로 변환하는 주파수/시간매핑부를 포함함이 바람직하다.

이하에서 첨부된 도면을 참조하여 본 발명의 바람직한 일실시예를 상세히 설명하기로 한다. 도 2는 본 발명에 의한 비트율 조절이 가능한 오디오 부호화장치의 바람직한 일실시예에 대한 구성을 블록도로 도시한 것으로서, 양자화처리부(230) 및 비트패킹부(240)로 이루어진다.

상기 양자화처리부(230)는 입력 오디오신호를 신호처리하여 소정의 부호화대역별로 양자화하는 블록으로서, 시간/주파수 매핑부(200), 심리음향부(210) 및 양자화부(220)로 이루어진다. 상기 시간/주파수 매핑부(200)는 시간영역의 입력오디오 신호를 주파수 영역의 신호로 변환한다. 시간상으로 인간이 인지하는 신호의 특성 차이가 그리 크지 않지만, 이렇게 변환된 주파수 영역의 신호들은 인간의 음향심리모델에 따라 각 대역에서 인간이 느낄 수 있는 신호와 느낄 수 없는 신호의 차이가 크기 때문에 각 주파수 대역에 따른 양자화 비트를 다르게 할당할 수 있음으로써 압축의 효율을 높일 수 있다.

심리음향부(210)는 상기 시간/주파수 매핑부(200)에 의해 주파수 성분으로 변환된 입력 오디오 신호들을 소정의 대역(subband) 신호들로 묶고, 각 신호들의 상호작용으로 인해 발생하는 마스킹현상을 이용하여 각 대역(subband)에서의 마스킹 문턱치(masking threshold)를 계산한다.

양자화부(220)는 각 대역의 양자화 잡음이 마스킹 문턱치보다 작아지도록 소정의 부호화 대역별로 양자화한다. 즉 인간이 들어도 느끼지 못도록 각 대역의 양자화 잡음의 크기가 상기 마스킹 문턱값보다 작도록 각 대역의 주파수 신호들을 스칼라 양자화를 사용한다. 음향심리부에서 계산한 마스킹문턱치와 각 대역에서 발생하는 잡음(noise)의 비율인 NMR(Noise-to-Mask Ratio)를 이용하여 전대역의 NMR 값이 0dB이하가 되도록 양자화를 행한다. NMR 값이 0dB이하라는 것은 양자화잡음에 비해 마스킹값이 높다는것을 나타내는데, 이것은 양자화잡음을 사람이 들을 수 없다는 의미이다.

상기 비트패킹부(240)는 비트율이 가장 낮은 기본계층에 상응하는 부가정보(side informatio) 및 양자화된 데이터들을 부호화하고, 상기 기본계층에 대한 부호화가 끝나면 그 다음 계층에 대한 부가정보 및 양자화된 데이터들을 부호화하여 이를 모든 계층에 대해 수행하여 비트스트림을 형성한다. 상기 비트패킹부(240)의 각 계층에서의 양자화된 데이터의 부호화는 상기 양자화된 데이터 각각을 소정의 동일 개수의 비트로 이루어지는 이진 데이터로 표현하여 비트단위로 나누고, 상기 분할된 비트들로부터 중요도가 가장 높은 최상위 비트로 이루어지는 최상위 비트시퀀스(sequence)부터 최하위 비트시퀀스 순서대로 소정의 부호화 방법에 의해 부호화한다. 그리고 상기 양자화된 데이터가 부호(sign)데이터와 크기(magnitude)데이터로 이루어질 때 상기 비트패킹부(240)는 상기 비트분할된 데이터들 중 중요도가 같은 동일 순위의 비트들에 대한 크기데이터를 모아서 부호화하고, 상기 부호화된 크기데이터 중 영이 아닌 크기데이터에 해당하는 부호(sign)데이터 중 부호화되지 않은 부호데이터를 부호화하고, 상기 크기데이터와 부호데이터의 부호화를 최상위 비트부터 하위비트로 내려가며 순차적으로 수행한다.

한편, 상기 부호화장치의 동작을 설명하기로 한다. 일단 입력된 오디오 신호는 부호화되면 비트스트림으로 만들어진다. 이를 위해 먼저 입력신호는 시간/주파수 매핑부(200)에서 MDCT(Modified Discrete Cosine Transform)이나 대역 분할 필터(subband filter)에 의해 주파수 영역의 신호로 변환된다. 그리고 심리음향부(210)는 상기 주파수 신호를 적당한 대역(subband)으로 묶어서 마스킹 문턱값을 구한다. 여기서 사용되는 대역은 주로 양자화 과정에 사용되기 때문에 양자화대역이라 부르게 된다. 양자화부(220)에서는 인간이 들어도 느끼지 못도록 각 양자화대역의 양자화 잡음의 크기가 마스킹 문턱값보다 작도록 스칼라 양자화 한다. 이런 조건을 만족하도록 양자화 과정을 거치고 나면, 각 대역에 대한 양자화 스텝 크기값과 양자화된 주파수 값들이 생성된다.

사람은 심리음향학적인 측면에서 낮은 주파수에서는 세밀한 간격의 주파수 성분이라도 쉽게 구분한다. 그러나 주파수가 높아질수록 사람이 구분할 수 있는 주파수 간격은 넓어진다. 그래서 도 5과 같이 낮은 주파수의 양자화대역은 대역폭(bandwidth)이 좁고, 높은 주파수의 양자화대역은 대역폭이 크다. 하지만, 부호화할 때는 부호화 과정의 용이성을 위해 도 5에 도시된 바와 같은 양자화 대역을 사용하지 않고, 대역 크기가 엇비슷한 부호화 대역을 사용한다. 즉 도 5에서처럼 대역 폭이 작은 경우에는 몇 개의 양자화대역을 묶어서 하나의 부호화대역이 되고, 대역 폭이 큰 경우에는 1개의 대역이 부호화 대역이 된다. 따라서, 전체적인 부호화 대역은 그 크기가 서로 엇비슷하도록 조절된다.

1. 데이터 중요도에 따른 부호화

우선 양자화된 값들의 부호(sign) 값을 따로 보관하고 절대값을 취해서 모든 양자화 값의 sign이 (+)가 되도록 한다. 각 부호화 대역내의 양자화된 주파수 값들 중에서 가장 큰 절대값을 가지는 값을 찾아 각 대역별로 신호를 표현하기 위해 필요한 양자화 비트수를 계산한다.

일반적으로 최상위비트(MSB, Most Significant Bit) 1비트의 중요도는 최하위비트(LSB, Least Significant Bit) 1비트의 중요도에 비해 훨씬 높다. 하지만, 기존의 방식으로 부호화를 할 경우, 이런 중요도에 관계없이 부호화가 되기 때문에 만약 전체 비트스트림 중 앞에서부터 일부의 비트스트림만을 사용해야 한다고 한다면 뒤쪽에 사용되지 못하는 비트스트림에 포함된 정보에 비해 덜 중요한 정보들이 앞쪽의 비트스트림에 상당히 많이 포함되어 있게 된다.

그래서 본 발명에서는 각 대역의 양자화된 신호들을 최상위 비트부터 최하위비트의 순서로 부호화하는 방식을 사용한다. 즉, 각 양자화된 신호들을 2진수로 표현하고 각 주파수 성분의 양자화된 값들을 비트 단위로 낮은 주파수 성분에서 높은 주파수 성분 순서로 처리하게 된다. 먼저, 각 주파수 성분의 최상위비트를 구해서 순서대로 1비트씩 모아서 부호화하고 난 후, 그 다음 상위비트를 부호화하고 순서대로 최하위비트까지 처리한다. 이런 순서로 부호화를 함으로써 좀 더 중요한 정보가 앞에서부터 비트스트림을 형성하도록 하는 것이다.

만일 각각 4비트로 8개의 양자화된 값이 다음과 같이 2진수로 나타내어 진다고 가정하자.

LSB MSB

0:1001

1:1000

2:0101

3:0010

4:0000

5:1000

6:0000

7:0100

기존의 방식대로 하면 우선 가장 낮은 주파수 성분의 값인 1001을 부호화하고 다음에 1000, 0101, 0010의 순으로 부호화한다. 하지만 본 발명의 방법에 의하면, 먼저 가장 낮은 주파수 성분의 최상위비트인 1과 다음에 0 1 0 0...등의 최상위비트 값을 구해서 순서대로 적당하게 몇 비트 씩 묶어서 처리한다. 예를 들어, 4비트씩 한 단위로 부호화를 한다면 1010을 부호화하고 다음에 0000의 순으로 부호화하고 최상위비트의 부호화가 끝나면 그 다음의 상위비트값들을 구해서 0001, 0000...의 순서로 부호화를 하고 최하위비트까지 같은 방식으로 부호화를 한다. 이 때 부호화하는 방법은 각 경우에 대한 적당한 확률 분포를 구해서 각 경우에 대한 호프만 부호화(huffman coding)나 산술 부호화(arithmetic coding) 등의 무손실 부호화 방식을 사용함으로써 효율적으로 압축하여 부호화를 할 수 있다.

2. 부호(sign) 비트를 포함한 부호화

양자화된 데이터의 sign 비트에 대한 부호화에 대해 설명하기로 한다. 통상적으로 상기 부호비트는 최상위비트에 위치한다. 따라서 최상위비트부터 부호화하면 부호비트가 가장 중요한 정보로 인식되어 먼저 부호화된다. 이럴 경우 비효율적인 부호화가 될 수 있다. 즉 최상위 비트부터 하위비트로 내려가며 부호화할 경우 처음 1이 나오기 전까지는 양자화된 값은 0으로 간주가 된다. 이 경우에는 sign 값이 무의미하다. 예를 들어, 양자화된 값이 2진수로 00011의 5비트로 표현될 때 상기 5비트 중 상위 3비트만이 부호화된다면 이를 복원하면 00000의 값으로 복원된다. 따라서, 이 값은 sign비트가 있다해도 아무런 쓸 데 없는 정보가 될 것이다. 하지만, 4비트가 부호화된다면 00010의 값이 됨으로써 sign값을 비로소 상당히 중요한 의미를 가지게 된다.

상위 비트에서 0이 나오다가 1이 처음나왔다는 것은 양자화된 값이 0이 아닌 어떤 값으로 복호화된다는 의미가 되므로 sign값이 중요한 의미를 가지게 된다. 각 주파수 성분들을 최상위비트부터 표현해가면서 처음으로 0이 아닌 1이 나오면 이 값 다음에는 다른 값들을 부호화하기 전에 (+)인지 (-)인지에 sign값을 부호화해야 한다. 예를 들면 앞의 예에서 최상위비트를 부호화할 때 먼저 1010을 부호화하고 나서 sign비트의 부호화가 필요한지에 대한 결정을 한다. 이 때 첫번째 주파수 성분과 3번째 주파수 성분에서 0이 아닌 값이 처음으로 부호화가 되었으므로, 상기 이 두 주파수 성분에 대한 sign 비트를 순서대로 부호화한다. 그 다음에 0000을 부호화한다. 만약, 최하위비트(LSB)들을 부호화할 때는 1100을 부호화하고 나서 sign비트의 필요여부를 결정한다. 이 경우에 1이 두 개가 나왔으나 첫번째 1은 앞서 최상위비트에서 1이 나온적이 있으므로 이미 sign비트가 부호화가 되었다. 따라서 sign비트를 부호화할 필요가 없고 두 번째 1은 상위비트들에서 1이 나온 적이 없으므로 sign비트를 부호화한다. 상기 sign 비트를 부호화하고 난 후, 다음 최하위 비트들인 0100을 부호화하게 된다.

3. 개선된 부호화 방법

상술한 부호화 방법을 적용할 경우 비트율이 낮을 경우에는 다음과 같이 부호화 순서를 다음과 같이 변형시키는 것이 효과적일 수 있다. 일반적으로 사람의 귀에 민감한 것은 각 주파수 성분에 대한 (+), (-) 윤곽이다. 여기서 제안하는 부호화 방법은 sign 비트가 부호화된 주파수 성분들에 대해서는 부호화를 뒤로 미루고 아직 sign 비트가 부호화되지 않아서 0으로 복원될 주파수 성분들에 대해서만 부호화를 진행한다. 이렇게 하여 sign 비트에 대한 부호화가 완료되면, 상기 뒤로 미룬 데이터들을 상술한 방법에 의해 부호화한다.

이를 상기 예를 이용하여 구체적으로 설명하면, 먼저 최상위 비트에서는 아직 sign값들이 부호화가 된 주파수 성분들이 없기 때문에 모두 부호화한다. 그리고 그 다음 상위 비트는 0001, 0000의 순서로 비트들이 나오지만 0001에서 1번째의 0과 3번째 0은 최상위 비트에서 sign비트가 부호화가 되었기 때문에 일단 부호화를 보류하고 2번째 0과 4번째 1만이 남으므로 01을 부호화한다. 그리고 여기서 나온 1은 상위 비트에서 1이 나온 적이 없으므로 이 주파수 성분에 대한 sign 비트를 부호화한다. 그리고 다음의 0000값을 보면 상위 비트들에서 sign 비트들을 부호화한 적이 없으므로 0000을 부호화하게 된다. 이렇게 최하위 비트까지의 sign 비트를 부호화하고 난 후 나머지 부호화되지 않은 정보들을 상위비트부터 순서대로 상술한 방법을 이용하여 부호화한다.

4. 스케일 조절이 가능한(scalable) 비트스트림의 구성

본 발명에서는 오디오 신호를 기본계층(base layer)과 여러개의 enhancement layer로 이루어지는 계층적인 비트스트림으로 부호화한다. 상기 기본계층은 비트율이 가장 낮은 계층이며, 상기 enhancement layer는 상기 기본계층보다 비트율이 높고, 상기 enhancement layer가 진행될수록 비트율은 더욱 높아진다.

상기 base layer의 앞부분에는 최상위비트들만 표현이 됨으로써 주파수 성분 전체에 대한 윤곽만이 부호화가 되다가 하위비트로 내려가면서 좀 더 많은 비트가 표현됨으로써 좀 더 세밀한 정보를 나타내게 된다. 그리고 비트율이 증가하면서 즉 계층이 진행됨(enhance)에 따라 세밀한 데이터값에 대한 부호화가 이루어지므로 양질의 음질을 나타낼 수 있다.

이렇게 표현된 데이터를 이용하여 scalable bitstream 을 구성하는 방법은 다음과 같다. 먼저, 기본계층(base layer)에 사용될 부가정보 중 각 양자화 대역의 양자화 비트 정보에 대한 부호화를 한다. 그리고 양자화된 값들에 대한 정보를 최상위 비트부터 최하위비트로, 그리고 낮은 주파수부터 높은 주파수의 순서대로 부호화한다. 물론 현재 부호화를 하고 있는 양자화 비트보다 어떤 대역의 양자화비트가 작을 경우에는 부호화를 하지 않고 지나치며, 그 양자화대역은 양자화비트가 그 대역의 양자화 비트와 같아질 때 비로소 부호화를 한다. 그리고 각 계층의 신호를 부호화할 때 아무런 대역 제한이 없으면 상당히 귀에 거슬리는 소리를 듣게 된다. 이는 최상위비트부터 하위비트 순서로 대역에 관계없이 부호화하면 비트율이 낮은 계층의 신호를 복원하면 신호가 있다가 없다가 하는 현상 때문이다. 따라서 비트율에 따라 적당하게 대역을 제한해서 사용하는 것이 좋다.

기본계층에 대한 부호화가 끝나면 그 다음 enhancement 계층에 대한 부가정보와 오디오데이터의 양자화값을 부호화한다. 이러한 방법으로 모든 계층의 데이터를 부호화한다. 이렇게 부호화된 모든 정보들을 모아서 비트스트림을 구성하게 된다.

이상에서, 상기 부호화장치에서 만들어진 비트스트림은 도 3에 도시된 바와 같이 비트율에 따라 하위 계층의 비트스트림이 상위 계층의 비트스트림에 포함되어 있는 계층 구조로 된다. 기존의 비트스트림들은 먼저 부가 정보들을 먼저 부호화하고 난 후 나머지 정보들을 부호화해서 비트스트림을 만들지만 본 발명에서는 도 3과 같이 각 계층에 필요한 부가 정보들은 각각 계층별로 나뉘어서 부호화 된다. 또한 기존의 부호화 방식에서는 양자화된 데이터를 모두 샘플 단위로 차례로 부호화를 했지만 본 발명에서는 양자화된 데이터를 2진데이터로 나타내고 2진데이터의 최상위비트(Most Significant Bit, MSB)에서부터 부호화를 해서 허용가능한 비트량내에서 비트스트림을 구성하는 방식을 사용한다.

한편, 상기 부호화장치의 동작을 구체적인 예를 들어 보다 상세히 설명하면 다음과 같다. 본 발명은 여러 계층의 비트율(bitrate)에 대한 정보를 하나의 비트스트림에서 도 3과 같이 계층적인 구조를 가지고 오디오 신호를 부호화함에 있어 중요한 신호 성분의 순으로 표현한다. 따라서 이렇게 만들어진 비트스트림은 사용자의 요구 또는 전송선로의 상태에 따라 가장 높은 비트율의 비트스트림에 포함된 낮은 비트율의 비트스트림들을 간단히 재구성해서 낮은 비트율의 비트스트림을 만들어 낼 수 있다. 즉, 실시간으로 부호화장치에서 만들어진 비트스트림이나 또는 어떤 매체에 저장되어 있는 비트스트림을 사용자의 요구에 따라 원하는 비트율에 대한 비트스트림으로 만들어 전송할 수 있다. 또한 사용자가 완전한 비트스트림을 가지고 있더라도 사용자의 하드웨어의 성능이 좋지 않든지 사용자가 복호화기의 복잡성을 낮추고자 한다면 이 비트스트림 중 일부만을 가지고 복원할 수 있게 함으로써 복잡성을 낮출 수 있으므로 복잡성 조절을 가능하게 할 수 있다.

한 예로, 기본계층(Base Layer)은 16kbps, 최상위 계층(Top Layer)은 64kbps이고, 각 계층(Enhancement Layer)을 8kbps 간격의 비트율을 가지는 비트스트림을 구성할 수 있다. 즉, 16, 24, 32, 40, 48, 56, 64kbps 7계층의 스케일 조절이 가능한 비트스트림을 구성하는 예이다. 따라서 부호화기에서 구성되는 비트스트림은 도 3과 같은 계층적 구조를 가지고 구성되기 때문에 최상위 계층인 64kbps에 대한 비트스트림내에 각 계층(16, 24, 32, 40, 48, 56, 64kbps)에 대한 비트스트림이 포함되어 있다. 만일 어떤 사용자가 최상위 계층에 대한 데이터를 요구하면 이 비트스트림을 어떤 가공도 없이 전달한다. 또 다른 사용자가 기본계층에(16bps에 해당하는) 대한 데이터를 요구하면 단순히 앞부분의 비트스트림만을 잘라내어 전달한다.

각 계층은 도 8와 같이 비트율에 따라서 대역폭이 제한되어 있어 비트율에 따라 끝 양자화 대역이 다르다. 입력 데이터는 48kHz로 샘플링된 PCM데이터이고, 1프레임의 크기는 1024개이다. 64 kbps의 비트율의 경우에 1프레임에서 쓸 수 있는 비트의 수는 평균 $64000\text{비트}/1\text{초} * (1024/48000) = 1365.3333\text{비트}$ 가 된다. 마찬가지로 각 비트율에 따라서 1프레임에 사용할 수 있는 비트의 크기를 계산할 수 있다. 이렇게 계산된 비트의 크기 도 7과 같다.

양자화를 하기 전에 먼저 입력데이터에서 음향심리모델(Psychoacoustic Model)을 이용해서 현재 처리되고 있는 프레임의 block type(long, start, short, stop)과 각 처리대역(processing band)의 SMR(Signal-to-Masked Threshold Ratio)값, short block인 경우 영역정보 그리고 음향심리모델과 시간/주파수의 동기를 맞추기 위해 시간지연된 PCM데이터들을 만들어 시간/주파수 매핑부에 전달한다. 음향심리모델을 계산하는 방법은 ISO/IEC 11172-3의 Model 2를 사용한다.

음향심리모델의 출력인 block type 에 따라 시간/주파수 맵핑부에서는 MDCT를 이용하여 시간영역의 데이터를 주파수 영역의 데이터로 변환한다. 이 때 block의 크기는 long/start/stop block인 경우 2048이고 short block 인 경우 크기가 256인 MDCT를 8번한다. 여기까지의 과정은 기존의 MPEG-2 NBC[13]에서 사용되는 것과 같은 방식을 사용한다.

주파수영역으로 변환된 데이터는 도 5와 같은 양자화 대역으로 주파수 성분들을 묶어서 양자화 대역의 SNR(Signal-to-Noise Ratio)값이 청각 심리모델의 출력값인 SMR값보다 작아지도록 stepsize 를 증가시켜 가면서 양자화한다. 양자화는 스칼라양자화(scala quantization)를 사용하며 기본적인 양자화 step size의 간격은 $2^{1/4}$ 를 사용한다. 양자화는 NMR값이 0dB 이하가 되도록 수행한다. 이때 얻어지는 출력은 양자화된 데이터와 각 처리대역의 양자화 step size에 대한 정보이다. 그리고 이렇게 양자화된 신호들을 부호화하기 위해서 먼저 이 양자화된 신호들을 부호화 대역별로 가장 큰 절대값을 찾아서 부호화에 필요한 가장 큰 양자화 비트를 계산한다.

비트스트림의 동기(synchronization) 신호를 위해 12 비트를 비트스트림에 넣어서 비트스트림의 시작에 대한 정보를 만든다. 그리고 나서 먼저 비트스트림 전체에 대한 크기를 부호화한다. 비트스트림의 부호화된 가장 높은 비트율에 대한 정보를 부호화한다. 이 정보가 있어야 낮은 비트율의 비트스트림으로 다시 만들고 이 비트율보다 높은 비트율을 요구했을 때 더 이상의 비트를 보내지 않을 수 있다. 그리고 다음으로 블럭타임을 부호화 한다. 그 다음의 부호화하는 과정은 블럭의 형태에 따라 약간 차이가 난다. 신호의 특성에 따라서 한 프레임의 입력 신호를 부호화하기 위해서 큰 블럭의 변환을 하는 경우와 8개의 짧은 블럭으로 나누어 변환을 하는 경우로 나뉜다. 이렇게 블럭의 크기가 바뀌기 때문에 부호화하는 과정에 약간의 차이가 있다.

먼저 긴 블럭의 경우에는 기본 계층의 대역폭은 4kHz이므로 처리대역은 12번째 양자화 대역까지이다. 먼저 12번째 양자화 대역까지의 기본계층의 비트할당정보를 부호화한다. 각 대역의 비트 할당정보로부터 최대 양자화 비트를 구하고 이 최대 양자화 비트값부터 상술한 방법에 의해 부호화를 한다. 그리고 차례로 다음 양자화된 비트들에 대한 부호화를 한다. 물론 현재 부호화를 하고 있는 양자화 비트보다 어떤 대역의 양자화 비트가 작을 경우 아무런 부호화를 하지 않고 지나치며 그 양자화대역은 양자화비트가 그 대역의 양자화 비트와 같아질 때 비로소 부호화를 한다. 그리고 이 때 처음으로 어느 대역을 부호화를 할 경우에 양자화 대역에 대한 stepsize 정보를 먼저 부호화를 하고 난 후 양자화된 주파수 성분들의 값중 그 양자화 비트에 해당되는 값들을 추출해서 부호화하게 된다. 최하위 계층의 비트율은 16kbps이므로 허용가능한 전체 비트는 336비트이다. 따라서 계속해서 사용되는 총 비트량을 계산하고 있다가 비트량이 336비트 이상이 되면 일단 부호화를 멈추게 된다. 양자화비트나 양자화스텝크기의 부호화 방법은 먼저 양자화 비트나 양자화 스텝크기들중에서 최소값과 최대값을 구하고 이 두 값의 차의 크기를 구하면 필요한 비트 수를 알 수 있다. 그래서, 실제로 이 부가 정보를 부호화하기 전에 먼저 최소값과 비트표현에 필요한 크기를 산술 부호화 방식에 의해 먼저 부호화해서 비트스트림에 저장해 놓고 나중에 실제로 부호화를 할 때는 최소값과의 부가정보의 차이를 부호화하는 방법을 사용한다.

마찬가지로 짧은 블록의 경우에는 크기가 긴 블록의 1/8인 8개의 블록으로 나누어 시간/주파수 맵핑과 양자화 과정을 거친 후 이 양자화된 데이터에 대한 무소실 부호화를 한다. 양자화 과정은 8개의 블록의 블록마다 각각 양자화를 하는 것이 아니고 심리음향부에서 8개의 블록을 3개의 영역으로 분리한 정보를 보내주는데 이 영역안에 있는 도 6와 같은 양자화 대역들을 모아서 긴 블록에서 하나의 대역처럼 처리한다. 그래서 3개 영역의 각 대역에 대한 양자화 스텝 크기 정보를 구하게 된다. 기본 계층의 대역폭을 긴 블록의 경우와 비슷하게 맞추기 위해 대역을 4번째 대역까지로 제한한다. 이 짧은 블록의 경우에는 8개의 작은 블록이 있으므로 도 6와 같이 1개의 블록에서 4개의 샘플단위로 부호화 대역으로 나누고 8개 블록의 부호화 대역을 묶어서 32개의 양자화된 신호들중에서 양자화 비트 정보를 구해서 사용한다. 먼저 제한된 대역내의 양자화비트 정보를 부호화한다. 대역제한된 성분내에서 최대 양자화 비트를 구하고 긴 블록과 마찬가지로 상술한 방법에 의해 부호화를 할 수 있다. 물론 현재 부호화를 하고 있는 양자화 비트보다 어떤 대역의 양자화비트가 작을 경우 아무런 부호화를 하지 않고 지나치며 그 양자화대역은 양자화비트가 그 대역의 양자화 비트와 같아질 때 비로소 부호화를 한다. 그리고 이 때, 이렇게 처음으로 어느 대역을 부호화를 할 경우에서 양자화 대역에 대한 stepsize 정보를 먼저 부호화를 하고 난 후 양자화된 주파수 성분들의 값중 그 양자화 비트에 해당되는 값들을 추출해서 부호화하게 된다.

기본 계층(16kbps)에 대한 비트스트림을 다 만들고 나면 그 다음계층(24kbps)에 대해 비트스트림을 작성한다. 24kbps에서의 대역폭은 8kHz이므로 19번째 대역까지의 주파수 성분들을 부호화한다. 19번째 대역까지의 부가정보는 이미 기록되어 있으므로 13번째부터 19번째 대역까지의 부가정보만을 기록한다. 최하위 계층에서 각 대역의 아직 부호화되지 않고 남아 있는 양자화 비트와 새로이 추가된 대역의 양자화 비트를 비교하여 최대 양자화 비트를 구하고 이 최대 양자화 비트부터 기본 계층에서와 같은 방식으로 부호화를 해 나가다가 사용된 총 비트량의 크기가 24 kbps에서 쓸 수 있는 비트 수보다 커지면 부호화 과정을 멈추고 다음 계층의 비트스트림을 작성할 준비한다. 이러한 방법으로 나머지 계층, 즉, 32, 40, 48, 56, 64kbps에 대한 비트스트림을 작성할 수 있다. 이렇게 구성된 비트 스트림은 도 3과 같은 구조로 구성된다.

한편, 상기 부호화장치에 의해 생성된 비트스트림을 복호화하는 복호화 장치를 상세히 설명하기로 한다. 도 4는 상기 복호화 장치의 구성을 블록도로 도시한 것으로서, 비트스트림분석부(400), 역양자화부(410) 및 주파수/시간 맵핑부(420)로 이루어진다.

상기 비트스트림 분석부(400)는 계층적 구조를 가진 비트스트림에서 상기 계층이 생성된 순서에 따라 양자화 스텝의 크기 및 각 대역에 할당된 양자화비트수를 포함하는 부가정보 및 양자화된 데이터를 복호화하되, 상기 비트스트림을 구성하고 있는 비트들의 중요도를 분석하여 중요도가 높은 비트부터 중요도가 낮은 비트로 내려가면서 상기 계층별로 복호화한다. 상기 역양자화부(410)는 복호화된 양자화 스텝의 크기와 양자화된 데이터들을 가지고 원래 크기의 신호로 복원한다. 상기 주파수/시간 맵핑부(420)는 주파수 영역의 오디오 신호를 다시 시간 영역의 신호로 변환해서 사용자가 재생할 수 있도록 해 준다.

상기 복호화장치의 동작을 설명하기로 한다. 상술한 부호화 장치에 의해 생성된 비트스트림의 복호화 과정은 부호화 과정의 역순을 거치면 된다. 먼저 기본 계층에 대한 정보를 복호화한다. 간단히 그 과정을 살펴보면, 먼저 기본계층의 부가정보 중 각 양자화 대역의 양자화 비트 정보에 대한 복호화를 한다. 이렇게 복호화된 양자화 비트들 가운데 가장 큰 값을 구한다. 상술한 부호화 순서에서와 마찬가지로 중요도가 가장 큰 비트부터 낮은 비트 순으로, 그리고 낮은 주파수부터 높은 주파수의 순서대로 비트스트림에서 양자화된 값들을 서서히 복호화한다. 물론 현재 복호화를 하고 있는 양자화 비트보다 어떤 대역의 양자화비트가 작을 경우 아무런 복호화 과정을 거치지 않고 지나치며 그 양자화대역은 양자화비트가 그 대역의 양자화 비트와 같아질 때 비로소 복호화를 시작한다. 그리고 이렇게 양자화된 값들에 대한 복호화가 진행되는 중간 중간에 이렇게 처음으로 어느 양자화 대역의 신호를 복호화를 할 경우에서 양자화 대역에 대한 stepsize 정보가 비트스트림에 저장되어 있기 때문에 이 정보를 먼저 복호화한 후 양자화된 값들에 대한 복호화를 계속하게 된다.

기본계층에 할당된 크기의 비트스트림에 대한 복호화가 끝나면 그 다음계층에 대한 부가정보와 오디오데이터의 양자화값을 복호화한다. 이러한 방법으로 모든 계층의 데이터를 복호화할 수 있다. 이렇게 복호화 과정을 거쳐 양자화된 데이터들은 앞서 부호화기와 반대의 순서로, 도 4에 도시된 역양자화부(410)와 주파수/시간 맵핑부(420)를 거쳐 복원된 신호를 만들 수 있다.

발명의 효과

본 발명은 다양한 사용자의 요구에 부응하기 위하여 비트스트림의 구성을 유연하게 한다. 즉, 사용자의 요구에 따라 여러계층의 비트율에 대한 정보를 중복없이 하나의 비트스트림에 결합시킴으로써 좋은 음질의 비트스트림을 제공할 수 있다. 또한, 송신단과 수신단 사이에 어떤 변환기도 필요없으며 전송선로의 상태나 사용자의 어떠한 요구도 수용할 수 있는 장점이 있다.

그리고 비트율조절이 가능한 형태의 비트스트림이므로 하나의 비트스트림안에 여러 개의 비트율의 비트스트림을 포함하고 있다. 또한 본 발명에서는 NMR이 0dB이하가 되도록 1번 양자화를 하고 난 후 더 이상의 비트조절부가 없기 때문에 부호화기의 복잡도가 낮다.

그리고 각 계층에서 앞 계층에서 양자화된 신호와 원래 신호의 차이를 다시 처리하여 부호화하지 않고 양자화 비트의 중요도에 따라서 부호화하기 때문에 부호화기의 복잡도가 낮다. 또한 하나의 비트스트림 안에 여러 개의 비트스트림을 포함하고 있기 때문에 여러계층의 비트스트림을 간단하게 생성할 수 있고 변호화기(transcoder)의 복잡도가 낮다.

그리고 전체 비트스트림에서 각 대역에 대한 부가정보를 한번만 사용함으로써 좋은 음질을 제공할 수 있다. 또한 비트율이 낮아지면 대역이 제한되어 있기 때문에 주로 부호화/복호화에서 복잡성의 대부분을 차지하는 필터의 복잡성이 상당히 감소하기 때문에 비트율에 비례해서 부호화기/복호화기의 실제 복잡성도 감소하게 된다. 또한 사용자의 복호화기의 성능, 전송선로의 대역폭/혼잡도 또는 사용자의 요구에 따라서 비트율이나 복잡성 조절이 가능하다.

(57) 청구의 범위

청구항 1.

오디오 신호를 기본계층(base layer)과 소정 개수의 향상계층(enhancement layer)으로 이루어지는 계층적인 데이터스트림으로 부호화하는 방법에 있어서, 입력 오디오 신호를 신호처리하여 소정의 부호화 대역별로 양자화하는 양자화처리단계, 상기 기본계층에 상응하는 양자화된 데이터들을 소정의 layer 크기 내에서 부호화하는 기본계층 부호화단계, 상기 부호화된 계층의 다음 enhancement 계층에 속하는 양자화된 데이터 및 상기 부호화된 계층에 속하는 부호화되지 않고 남아있는 양자화된 데이터들을 소정의 계층 크기 내에서 부호화하는 계층부호화 단계 및 상기 계층부호화단계를 모든 계층에 대해 수행하는 순차부호화단계를 포함함을 특징으로 하고, 상기 기본계층부호화단계, 계층부호화단계 및 순차 부호화단계의 부호화는 부호화하고자 하는 해당 계층에 상응하는 양자화된 데이터를 소정의 동일 개수의 디지털로 표현하는 제1단계 및 상기 동일 개수의 디지털로 표현된 양자화된 데이터들의 중요도가 가장 높은 디지털들로 이루어지는 최상위 디지털 시퀀스부터 중요도가 낮은 디지털 시퀀스 순으로 소정의 부호화 방법으로 부호화하는 제2단계로 이루어지는, 비트율 조절이 가능한 오디오 부호화 방법.

청구항 2.

제1항에 있어서, 상기 제1단계 및 제2단계의 수행은 낮은 주파수성분부터 높은 주파수성분 순서대로 수행함을 특징으로 하는, 비트율 조절이 가능한 오디오 부호화 방법.

청구항 3.

제1항에 있어서, 상기 기본계층부호화단계, 계층부호화단계 및 순차부호화단계의 부호화는 적어도 스텝크기(step size) 및 각 대역에 할당된 양자화비트수를 포함하는 부가정보를 소정의 부호화 방법을 사용하여 부호화함을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 방법.

청구항 4.

제1항 또는 제3항에 있어서, 상기 제1단계 및 제2단계의 디지털은 비트임을 특징으로 하는, 비트율 조절이 가능한 오디오 부호화 방법.

청구항 5.

제4항에 있어서, 상기 제2단계의 부호화는 상기 비트 시퀀스를 구성하고 있는 비트들을 소정 개수의 비트 단위로 묶어서 부호화함을 특징으로 하는, 비트율 조절이 가능한 오디오 부호화 방법.

청구항 6.

제4항에 있어서, 상기 소정의 부호화방법은 무손실 부호화임을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 방법.

청구항 7.

제5항에 있어서, 상기 소정의 부호화방법은 무손실 부호화임을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 방법.

청구항 8.

제6항 또는 제7항에 있어서, 상기 무손실 부호화는 허프만 부호화(Huffman coding)임을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 방법.

청구항 9.

제6항 또는 제7항에 있어서, 상기 무손실 부호화는 산술 부호화(Arithmetic coding)임을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 방법.

청구항 10.

제1항에 있어서, 상기 양자화된 데이터가 부호(sign)데이터와 크기(magnitude)데이터로 이루어질 때, 상기 제2단계는 상기 동일 개수의 디지털로 표현된 양자화된 데이터들의 크기 데이터들의 중요도가 가장 높은 최상위 디지털(Most Significant Digit)들로 이루어지는 최상위 디지털 시퀀스(sequence)를 소정의 부호화방법에 의해 부호화하는 단계, 상기 부호화된 최상위 디지털 시퀀스 중 영이 아닌 데이터에 해당하는 부호 데이터를 부호화하는 단계, 상기 디지털 데이터의 부호화되지 않은 크기데이터 중 중요도가 가장 높은 디지털 시퀀스를 소정의 부호화 방법에 의해 부호화하는 크기부호화단계, 상기 제3단계에서 부호화된 디지털 시퀀스 중 영이 아닌 크기 데이터에 해당하는 부호데이터들 중 부호화되지 않은 부호(sign)데이터를 부호화하는 부호(sign)부호화단계 및 상기 크기부호화단계 및 부호부호화단계를 상기 디지털 데이터의 각 디지털에 대해 수행하는 단계를 포함함을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 방법.

청구항 11.

제10항에 있어서, 상기 제1단계는 상기 디지털 데이터들을 동일 개수의 비트를 가진 이진 데이터로 표현하는 단계이고, 상기 디지털은 비트임을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 방법.

청구항 12.

제10항에 있어서, 상기 부호화는 상기 크기 데이터 및 부호(sign) 데이터에 대한 각 비트 시퀀스를 구성하고 있는 비트들을 소정 개수의 비트 단위로 묶어서 부호화함을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 방법.

청구항 13.

제11항 또는 제12항에 있어서, 상기 소정의 부호화방법은 산술 부호화(Arithmetic coding)임을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 방법.

청구항 14.

제10항에 있어서, 상기 기본계층부호화단계, 계층부호화단계 및 순차부호화단계의 부호화는 적어도 스텝크기(step size) 및 각 대역에 할당된 양자화비트수를 포함하는 부가정보를 소정의 부호화 방법을 사용하여 부호화함을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 방법.

청구항 15.

제1항 또는 제10항에 있어서, 상기 양자화처리단계는 시간 영역의 입력 오디오 신호를 주파수 영역의 신호로 변환하는 단계, 상기 시간/주파수 맵핑에 의해 주파수 성분으로 변환된 입력 오디오 신호들을 소정의 대역(subband) 신호들로 묶고, 상기 각 대역(subband)에서의 마스킹 문턱치(masking threshold)를 계산하는 단계 및 각 대역의 양자화 잡음이 마스킹 문턱치보다 작아지도록 소정의 부호화 대역별로 양자화하는 단계로 이루어짐을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 방법.

청구항 16.

오디오 신호를 소정 개수의 다른 비트율로 이루어지는 계층적인 비트율을 갖도록 부호화하는 장치에 있어서, 입력 오디오 신호를 신호처리하여 소정의 부호화 대역별로 양자화하는 양자화처리부 및 기본계층에 상응하는 부가정보(side info) 및 양자화된 데이터들을 부호화하고, 상기 기본계층에 대한 부호화가 끝나면 그 다음 계층에 대한 부가정보 및 양자화된 데이터들을 부호하여 이를 모든 계층에 대해 수행하여 비트스트림을 형성하는 비트패킹부를 포함함을 특징으로 하고, 상기 비트패킹부의 각 계층에서의 양자화된 데이터의 부호화는 상기 양자화된 데이터 각각을 소정의 동일 개수의 비트로 이루어지는 이진 데이터로 표현하여 비트단위로 나누고, 상기 분할된 비트들로부터 중요도가 가장 높은 최상위 비트로 이루어지는 최상위 비트시퀀스부터 최하위 비트시퀀스 순서대로 소정의 부호화 방법에 의해 부호화하는, 비트율 조절이 가능한 오디오 부호화 장치.

청구항 17.

제16항에 있어서, 상기 디지털 데이터가 부호(sign) 데이터와 크기(magnitude) 데이터로 이루어질 때 상기 비트패킹부는 상기 비트분할된 데이터들 중 중요도가 같은 동일 순위의 비트들에 대한 크기데이터를 모아서 부호화하고, 상기 부호화된 크기데이터 중 영이 아닌 크기데이터에 해당하는 부호(sign)데이터 중 부호화되지 않은 부호데이터를 부호화하되, 상기 크기데이터와 부호데이터의 부호화를 최상위 비트부터 하위비트로 내려가며 순차적으로 수행함을 특징으로 하는, 비트율 조절이 가능한 오디오 부호화 장치.

청구항 18.

제16항 또는 제17항에 있어서, 상기 비트패킹부는 중요도에 따라 비트들을 모아서 부호화할 때, 소정 개수의 비트단위로 묶어서 부호화함을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 장치.

청구항 19.

제16항 또는 제17항에 있어서, 상기 부호화 방법은 허프만 부호화 및 산술 부호화 중 어느 하나의 방식에 의해 부호화함을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 장치.

청구항 20.

제18항에 있어서, 상기 비트패킹부의 부호화는 낮은 주파수 성분부터 높은 주파수 성분 순서로 수행함을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 장치.

청구항 21.

제16항 또는 제17항에 있어서, 상기 양자화처리부는 시간 영역의 입력 오디오 신호를 주파수 영역의 신호로 변환하는 시간/주파수 맵핑부, 상기 시간/주파수 맵핑에 의해 주파수 성분으로 변환된 입력 오디오 신호들을 소정의 대역(subband) 신호들로 묶고, 상기 각 대역(subband)에서의 마스킹 문턱치(masking threshold)를 계산하는 심리음향부 및 각 대역의 양자화 잡음이 마스킹 문턱치보다 작아지도록 소정의 부호화 대역별로 양자화하는 양자화부를 포함함을 특징으로 하는 비트율 조절이 가능한 오디오 부호화 장치.

청구항 22.

계층적인 비트율을 갖도록 부호화된 오디오 데이터를 복호화하는 방법에 있어서, 계층적 구조를 가진 데이터스트림에서 상기 계층이 생성된 순서에 따라 적어도 양자화 스텝 크기 및 각 대역에 할당된 양자화비트수를 포함하는 부가정보 및 양자화된 데이터를 복호화하되, 상기 데이터스트림을 구성하고 있는 비트들의 중요도를 분석하여 중요도가 높은 디지털부터 중요도가 낮은 디지털로 내려가면서 복호화하는 복호화단계, 상기 복호화된 양자화 스텝 크기와 양자화된 데이터들을 원래 크기의 신호로 복원하는 단계 및 상기 역양자화된 주파수 영역의 오디오 신호를 시간 영역의 신호로 변환하는 단계를 포함함을 특징으로 하는, 비트율 조절가능한 오디오 데이터 복호화 방법.

청구항 23.

제22항에 있어서, 상기 복호화단계의 디지털은 비트이고, 상기 복호화단계의 데이터 스트림은 비트스트림임을 특징으로 하는 비트율 조절가능한 오디오 데이터 복호화 방법.

청구항 24.

제23항에 있어서, 상기 복호화단계의 중요도에 따른 복호화는 소정의 개수의 비트로 이루어진 벡터단위로 복호화함을 특징으로 하는 비트율 조절가능한 오디오 데이터 복호화 방법.

청구항 25.

제23항 또는 제24항에 있어서, 상기 양자화된 데이터가 부호(sign)데이터 및 크기(magnitude)데이터로 이루어질 때, 상기 복호화단계의 복호화는 계층적 구조를 가진 데이터스트림에서 상기 계층이 생성된 순서에 따라 적어도 양자화 스텝의 크기 및 각 대역에 할당된 양자화비트수를 포함하는 부가정보 및 양자화된 크기 데이터를 복호화하되, 상기 데이터스트림을 구성하고 있는 비트들의 중요도를 분석하여 중요도가 높은 디지털부터 중요도가 낮은 디지털로 내려가면서 복호화하는 단계 및 상기 양자화된 데이터의 부호(sign)데이터를 복호화하여 이를 상기 복호화된 크기 데이터와 결합하는 단계로 이루어짐을 특징으로 하는 비트율 조절가능한 오디오 데이터 복호화 방법.

청구항 26.

제23항에 있어서, 상기 복호화단계에서 복호화는 산술복호화 방법에 의해 복호화함을 특징으로 하는 비트율 조절가능한 오디오 데이터 복호화 방법.

청구항 27.

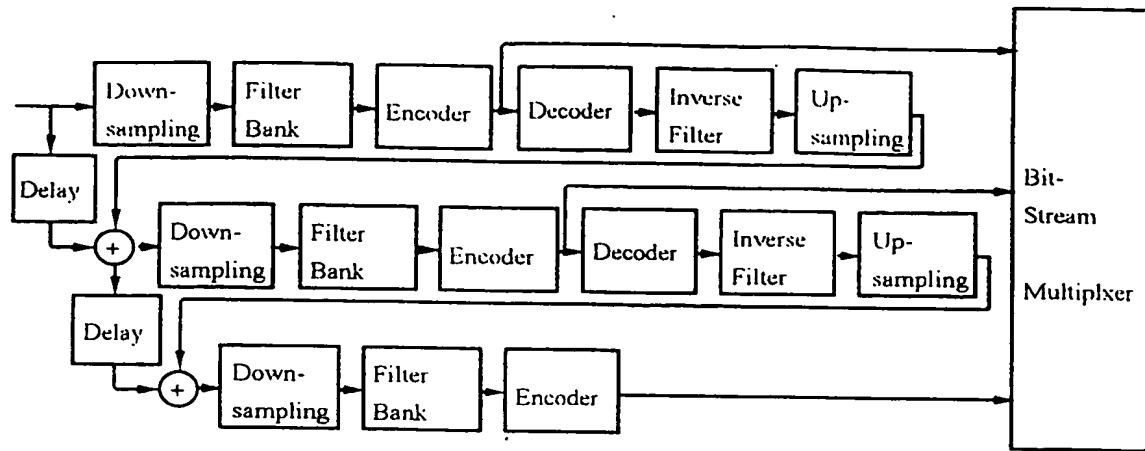
제23항에 있어서, 상기 복호화단계에서 복호화는 허프만복호화 방법에 의해 복호화함을 특징으로 하는 비트율 조절가능한 오디오 데이터 복호화 방법.

청구항 28.

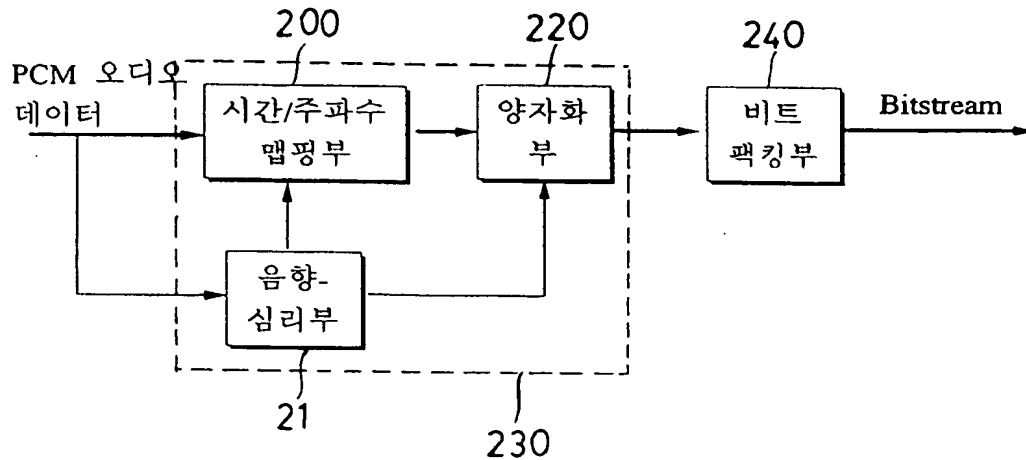
계층적인 비트율을 갖도록 부호화된 오디오 데이터를 복호화하는 장치에 있어서, 계층적 구조를 가진 비트스트림에서 상기 계층이 생성된 순서에 따라 적어도 양자화 스텝의 크기 및 각 대역에 할당된 양자화비트수를 포함하는 부가정보 및 양자화된 데이터를 복호화하되, 상기 비트스트림을 구성하고 있는 비트들의 중요도를 분석하여 중요도가 높은 비트부터 중요도가 낮은 비트로 내려가면서 상기 계층별로 복호화하는 비트스트림분석부, 상기 비트스트림분석부에서 복호화된 양자화 스텝 크기와 양자화된 데이터들을 원래 크기의 신호로 복원하는 역양자화부 및 상기 역양자화부에서 역양자화된 주파수 영역의 오디오 신호를 시간 영역의 신호로 변환하는 주파수/시간매핑부를 포함함을 특징으로 하는, 비트율 조절가능한 오디오 데이터 복호화 장치.

도면

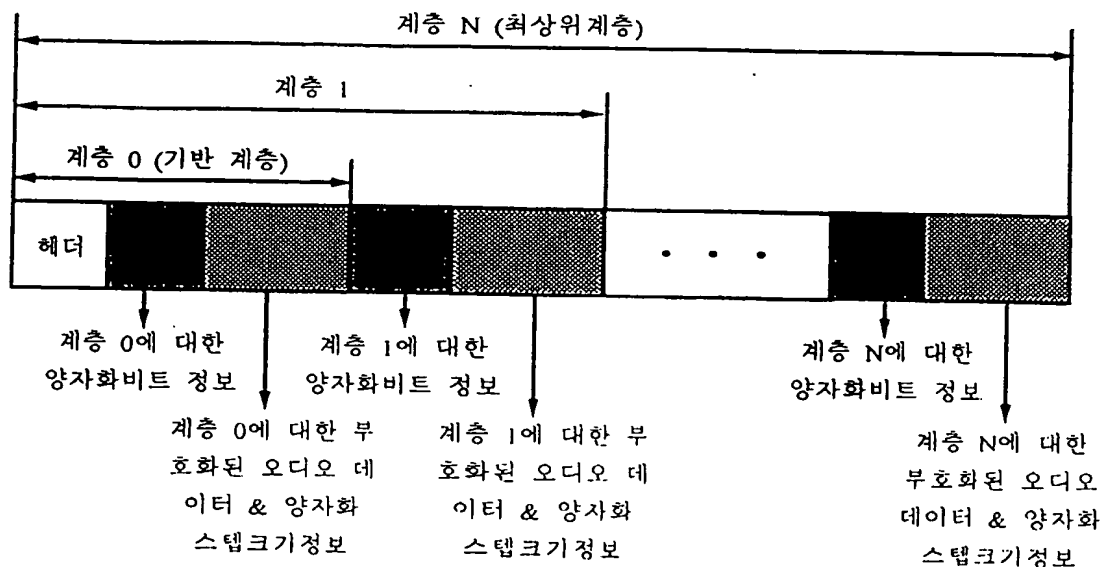
도면 1



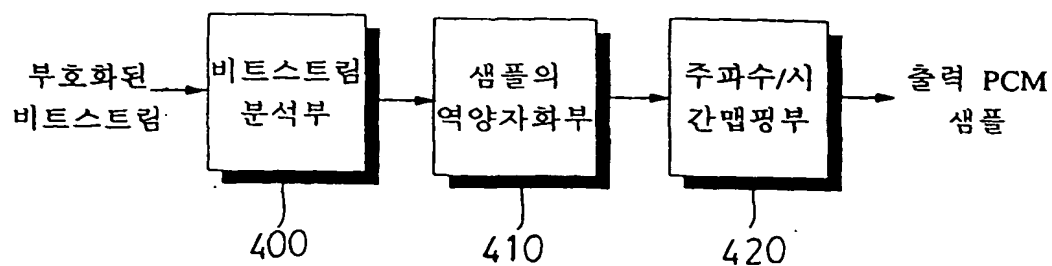
도면 2



도면 3



도면 4



도면 5

양자화 대역	부호화 대역	시각 index	문 index
0	0	0	7
1		8	15
2		16	23
3		24	35
4	1	36	47
5		48	59
6		60	71
7		72	83
8	2	84	99
9		100	115
10		116	131
11		132	147
12	3	148	163
13		164	195
14		196	227
15		228	259
16	4	260	291
17		292	323
18		324	354
19		356	387
20	5	388	419
21		420	451
22		452	483
23		484	515
24	6	516	555
25		556	599
26		600	634
27		644	687

도면 6

부호화 대역	양자화 대역	시작 index	끝 index
0	0	0	3
1	1	4	7
2	2	8	11
3	3	12	15
4	4	16	19
5	5	20	23
6	6	24	27
7		28	31
8	7	32	35
9		36	39
10	8	40	43
11		44	47
12	9	48	51
13		52	55
14		56	59
15	10	60	63
16		64	67
17		68	71
18	11	72	75
19		76	79
20		80	83
21		84	87

도면 7

비트율 (kbps)	16	24	32	40	48	56	64
비트수/ 프레임	336	512	680	848	1024	1192	1365

도면 8

비트율(kbps)	16	24	32	40	48	56	64
제한된 대역 (긴 블록)	0 - 12	0 - 19	0 - 21	0 - 23	0 - 25	0 - 27	0 - 27
제한된 대역 (짧은 블록)	0 - 4	0 - 7	0 - 8	0 - 9	0 - 10	0 - 11	0 - 11
대역폭	4 kHz	8 KHz	10 kHz	12 kHz	14 kHz	16 kHz	16 kHz

도면 9

영역 index	1번째 영역의 시작 블록 index	2번째 영역의 시작 블록 index	3번째 영역의 시작 블록 index
0	1	1	6
1	1	2	5
2	1	3	4
3	1	4	3
4	1	5	2
5	1	6	1
6	2	1	5
7	2	2	4
8	2	3	3
9	2	4	2
10	2	5	1
11	3	1	4
12	3	2	3
13	3	3	2
14	3	4	1
15	4	1	3
16	4	2	2
17	4	3	1
18	5	1	2
19	5	2	1
20	6	1	1